Published OnlineFirst January 4, 2019; DOI: 10.1158/1541-7786.MCR-18-1178

Cancer "-omics"

Molecular Cancer Research

Multiclonality and Marked Branched Evolution of Low-Grade Endometrioid Endometrial Carcinoma

Lorena Lazo de la Vega[1], Mia C. Samaha[1], Kevin Hu[2], Nolan R. Bick[1], Javed Siddiqui[1,3], Daniel H. Hovelson[2,3], Chia-Jen Liu[1,3], Cody S. Carter[1], Kathleen R. Cho[1], Andrew P. Sciallis[1], and Scott A. Tomlins[1,3,4,5]

Abstract

The molecular events driving low-grade endometrioid endometrial carcinoma (LGEC) development—like in many cancers—are incompletely understood. Hence, here we performed multiregion, comprehensive somatic molecular profiling of routinely processed formalin-fixed, paraffin-embedded (FFPE) material from 13 cases of LGEC totaling 64 minute, spatially defined cell populations ranging from presumed precursor lesions through invasive LGEC. Shared driving *PTEN*, *PIK3R1*, or *PIK3CA* mutations support clonal origin of the samples in each case, except for two cases with two clonally distinct neoplastic populations, consistent with unexpected multiclonality in LGEC development. Although substantial heterogeneity in driving somatic alterations was present across populations in nearly all cases, these alterations were usually clonal in a given population, supporting continued selection and clonal sweeping of driving alterations in populations with both precursor and LGEC histology. Importantly, *CTNNB1* mutational status, which has been proposed as both prognostic and predictive in LGEC, was frequently heterogeneous and subclonal, occurring both exclusively in precursor or cancer populations in different cases. Whole-transcriptome profiling of coisolated RNA from 12 lesions (from 5 cases) was robust and confirmed histologic and molecular heterogeneity, including activated Wnt signaling in *CTNNB1*-mutant versus wild-type populations. Taken together, we demonstrate clinically relevant multiclonality and intratumoral heterogeneity during LGEC development with important implications for diagnosis, prognosis, and therapeutic prediction. More broadly, our methodology is broadly scalable to enable high-throughput genomic and transcriptomic characterization of precursor and invasive cancer populations from routine FFPE specimens.

Implications: Multiregion profiling of LGEC populations using a highly scalable approach demonstrates clinically relevant multiclonality and intratumoral heterogeneity.

Introduction

Integrated genomic characterization of endometrial carcinoma (EC) by The Cancer Genome Atlas (TCGA) defined four groups based on histology, copy-number alterations (CNA), and mutations: *POLE* (ultramutated), microsatellite instability (hypermutated), CNA-high (serous-like), and CNA-low (endometrioid), consistent with clinical/pathologic/molecular endometrial carcinoma classification as type I [usually low-grade, endometrioid (LGEC)] and type II (high-grade, nonendometrioid; refs. 1, 2). Endometrioid endometrial carcinomas are thought to develop through hyperplastic precursor lesions characterized by architectural and nuclear atypia. Although criteria differ, systems based on (1) nuclear atypia and glandular complexity [World Health Organization (WHO)] or (2) molecular genetics/morphology (endometrial intraepithelial neoplasia) are widely used (3). Lesions classified by the first as atypical hyperplasia (AH)—more specifically complex atypical hyperplasia (CAH) when glandular complexity is present—and by the second as endometrial intraepithelial neoplasia (EIN) are now considered similar premalignant processes, and the terms are used interchangeably in the latest WHO classification system (2, 3). Endometrial hyperplasia without atypia, sometimes referred to as complex hyperplasia (CH), is thought to result from unopposed estrogen stimulation and has a lower risk of progression to LGEC than EIN/AH. LGEC and its precursors often display foci of squamous differentiation, a feature that is not typically seen in other types of endometrial carcinomas, such as serous or clear cell carcinomas. LGECs are usually CNA-low, non-hyper/ultramutated, lack *TP53* mutations, and frequently harbor somatic alterations affecting the PI3K, RTK/RAS, and Wnt signaling pathways (including recurrent mutations in *PTEN*, *PIK3R1*, *PIK3CA*, *KRAS*, and *CTNNB1*; ref. 1).

As reflected in calls to generate a Pre-Cancer Genome Atlas (PCGA), the molecular progression of EIN/AH to endometrial carcinoma, like in many cancers, is incompletely understood in part due to the technical challenges of profiling minute lesions/areas of interest often available only in routinely processed

[1]Department of Pathology, University of Michigan Medical School, Ann Arbor, Michigan. [2]Department of Computational Medicine and Bioinformatics, University of Michigan Medical School, Ann Arbor, Michigan. [3]Michigan Center for Translational Pathology, University of Michigan Medical School, Ann Arbor, Michigan. [4]Rogel Cancer Center, University of Michigan Medical School, Ann Arbor, Michigan. [5]Department of Urology, University of Michigan Medical School, Ann Arbor, Michigan.

Note: Supplementary data for this article are available at Molecular Cancer Research Online (http://mcr.aacrjournals.org/).

Corresponding Author: Scott A. Tomlins, University of Michigan Medical School, 1500 East Medical Center Drive, 7322 Rogel Cancer Center, Ann Arbor, MI 48109. Phone: 734-764-1549; Fax: 734-647-7950; E-mail: tomlinss@med.umich.edu

doi: 10.1158/1541-7786.MCR-18-1178

Published OnlineFirst January 4, 2019; DOI: 10.1158/1541-7786.MCR-18-1178

formalin-fixed paraffin-embedded (FFPE) specimens (4). Driving *PTEN* mutations occur early in type I endometrial carcinomas because they are generally found to coexist with other commonly mutated genes and were critical in defining EIN (5). However, whether EIN/AH usually progresses to endometrial carcinoma via linear versus branched evolution is unresolved. Limited intratumoral heterogeneity with respect to integrative classification of endometrial carcinomas has been reported, including 96% concordance of *CTNNB1* mutational status (6), and a next-generation sequencing (NGS)–based study of three pairs of EIN/AH and CNA-low LGEC supported clonal origin in all cases (7). In contrast, substantial mutational heterogeneity, supporting branched evolution, was reported in a study of 6 cases of matched, but spatially distinct EIN/AH and CNA-low LGEC (7), as well as in a hotspot NGS-based study of endometrial carcinoma from paired uterine aspirates and multiple regions at hysterectomy (8).

Understanding intratumoral heterogeneity in LGEC development is critical for the development of prognostic biomarkers. Although most patients with LGEC are cured by surgery alone, those that recur do poorly, arguing for the identification of prognostic biomarkers. Recently, Liu and colleagues and Kurnit and colleagues both reported that *CTNNB1* mutations were prognostic for shorter recurrence-free survival in patients with low-stage LGEC (9, 10). We were intrigued by this finding, as we had previously observed different *CTNNB1* mutations in paired primary uterine endometrial carcinoma (p.S45P) and tubal metastasis (p.S45F) components of a clinically type I high-grade endometrioid carcinoma that had a shared *PTEN* (p.R130X) mutation in both components (11). Likewise, we recently observed discordant *CTNNB1* mutations in the different components of a uterine endometrial carcinoma that had areas of conventional histology as well as areas with variant histology referred to as "corded and hyalinized" endometrial carcinoma (CHEC; p.G34E and p.S33C; C.S. Carter; in preparation). Hence, to comprehensively assess intratumoral heterogeneity in LGEC development, we performed multiregion profiling of matched spatially defined EIN/AH and LGEC components from routinely processed FFPE tissue specimens using a highly scalable, comprehensive multiplexed PCR-based NGS approach.

Materials and Methods

Cohort

With Institutional Review Board approval, we retrospectively identified patients with LGEC (FIGO grade I/II) at hysterectomy using a previously described electronic medical record search engine (12). We collected 14 cases with available archived FFPE tissues that had spatially and histologically distinct foci of both precursor (EIN/AH) and endometrial carcinoma. For each case, regions of interest were identified on hematoxylin and eosin (H&E)–stained slides and classified according to the WHO histologic system by board-certified pathologists (A.P. Sciallis and S.A. Tomlins) as CH, CAH, frankly invasive endometrial carcinoma, or frankly invasive endometrial carcinoma with squamous differentiation (ECsq). Regions were punched (1–3 punches) from the FFPE block using 21-gauge dispensing tips (0.510 mm inner diameter) followed by examination of an H&E recut to confirm localization. DNA and RNA from each punch were coisolated using the Qiagen Allprep FFPE DNA/RNA Kit (Qiagen) and quantified using the Qubit 2.0 fluorometer (Life Technologies) as described (13).

DNA NGS

We performed targeted, multiplexed PCR-based DNA NGS essentially as described (13) using panels targeting >130 cancer-related genes, including those recurrently mutated in LGEC (1), as described in detail in the Supplementary Methods. We used 20 to 24 ng of DNA per sample for library construction using the Ion Ampliseq library kit 2.0 (Life Technologies) with barcode incorporation and sequencing on the Ion Torrent Proton sequencer as described (13) and detailed in the Supplementary Methods. Data analysis was performed essentially as described to identify high-confidence, prioritized somatic mutations and CNAs using validated pipelines based on Torrent Suite 5.0.4.0 (11, 13, 14). All high-confidence somatic variants were visualized in Integrative Genomics Viewer (IGV), with selected validation by Sanger sequencing (Supplementary Methods and Supplementary Table S1). *POLE* hotspot mutation status was assessed by Sanger sequencing as they are not targeted by our panels. High-confidence somatic variants occurring at hotspots (>3 observations at that residue in COSMIC) in oncogenes, inframe indels in oncogenes or tumor-suppressor genes, or hotspot or deleterious (nonsense/frameshift/splice site altering variants) in tumor-suppressor genes were considered driving variants (11, 13). Case identity was confirmed in all populations by assessment of rare high-confidence SNPs.

Phylogenetic analysis

We conducted evolutionary analysis using PHYLIP v 3.695. For each tumor sample, the status of nonsynonymous somatic mutations were considered as characteristics for this analysis. Evolutionary trees were constructed using Dollop (Dollo and polymorphism parsimony methods) using polymorphism parsimony with default parameters.

Amplicon-based whole-transcriptome sequencing

We performed amplicon-based whole-transcriptome sequencing using the Ion Ampliseq Transcriptome Human Gene Expression Kit (Life Technologies) according to the manufacturer's instructions with 17.5 ng of RNA per sample, allowing for interrogation of ∼21,000 RNA transcripts. Library preparation was performed according to the manufacturer's instructions and as described above for DNA sequencing. Technical replicate libraries and templates were independently constructed and sequenced on separate chips. Reads were mapped and quantified using version 5.0.4.0 of TorrentSuite's (Life Science Technologies) coverageAnalysis plugin with the uniquely mapped reads option and default parameters. As described in detail in the Supplementary Methods, end-to-end reads were used for differential gene expression analysis using the R package edgeR (15, 16). Volcano plots were made using the R-package ggplot, and multiplicity was corrected by calculating a Q-value using Benjamini and Hochberg's FDR (17). Analyses performed to assess differentially expressed genes in relevant comparisons are described in the Supplementary Methods.

Immunohistochemistry

Polyclonal rabbit anti-amylase (AMY1A) primary antibodies [HPA045399 (562), 1:800; HPA045394 (560), 1:2,000] were selected based on confirmation of expression in expected tissues (pancreas and salivary glands) in the Human Protein Tissue Atlas (18). IHC staining was performed on 4 to 5 μm unstained FFPE slides using an automated protocol on the Ventana

Published OnlineFirst January 4, 2019; DOI: 10.1158/1541-7786.MCR-18-1178

Benchmark XT System using UltraView Universal DAB Detection Kit (Cat no. 760–500, Ventana Medical Systems). Staining was optimized and confirmed to show expected staining in pancreas and salivary gland tissues prior to staining endometrial carcinoma samples.

Results

Comprehensive genomic profiling of LGEC development

To assess the molecular landscape of LGEC development, we performed comprehensive DNA- and RNA-based NGS of 14 cases of FIGO grade 1 or 2 endometrial carcinoma with spatially defined minute precursors and/or endometrial carcinoma components using a highly scalable approach optimized for routine FFPE material (Supplementary Table S2 and Fig. 1). We obtained between 130 and 1,850 ng of extracted DNA (mean 987 ng), consistent with tens of thousands of cells from the punched regions. To identify oncogenic and tumor-suppressive somatic mutations and CNAs, we performed multiplexed PCR-based DNA NGS (mxDNAseq) on 70 spatially defined, minute ($\sim$1–2mm^2 surface area) cell populations (Supplementary Table S3) using panels targeting $\geq$130 genes, including essentially all recurrently altered genes in LGEC using extensively validated approaches (details regarding quality control of the sequencing can be found in the Supplementary Results and Supplementary Table S4). As described below, to validate the impact of observed histologic and somatic mutational heterogeneity, we also performed multiplexed PCR-based transcriptome NGS (mxRNAseq) on coisolated RNA from 12 cell populations.

As described in the Supplementary Results and shown in Supplementary Fig. S1, out of the 14 cases, 1 was classified as ultramutated. In the remaining 13 cases, no high-level, focal somatic CNAs were identified in any cell populations (Supplementary Fig. S2); hence, these cases were considered CNA-low LGEC. After exclusion of populations failing QC metrics, our cohort represented the full spectrum of LGEC development, including 2, 23, 27, and 12 populations ($n = 64$ total) classified as CH, CAH, invasive endometrial carcinoma, or invasive ECsq, respectively, as represented in Supplementary Fig. S3.

Across the 13 CNA-low LGEC cases, all cell populations harbored at least one clear driving somatic mutation in *PTEN*, *PIK3R1*, or *PIK3CA* (Fig. 2) consistent with the nearly universal deregulation of this pathway as a driving event in LGEC. Ten of 13 cases had at least one driving *PTEN* mutation detected in all cell populations (4 and 3 cases also had *PIK3CA* and *PIK3R1* mutations, respectively, in all cell populations) consistent with prior single gene and TCGA studies (1, 19–21). In the remaining three cases, no cell populations harbored *PTEN* mutations, but two invasive EC/ECsq cases harbored somatic driving clonal *PIK3R1* mutations (Cases 6 and 9), and one CAH case had somatic driving clonal *PIK3CA* mutations (Case 7).

We also identified recurrent, driving mutations across our LGEC cases in *CTNNB1*, *FBXW7*, *KRAS*, and *FGFR2*, consistent with bulk sequencing of LGEC (1). Importantly, across the 64 populations, we identified an average of 4 (range, 2–5; Supplementary Table S5) driving somatic mutations in the above-described seven genes, making LGEC an ideal system to assess clonality and intratumoral heterogeneity using a very limited subset of the genome. Of note, no significant difference in the number of prioritized mutations was observed between CH/CAH and EC/ECsq populations (average 3.0 vs. 3.3, two-tailed unpaired t test, $P = 0.18$).

Multiclonality in LGEC development

Although 11 of 13 LGEC cases were clonal based on shared *PTEN*, *PIK3R1*, or *PIK3CA* mutations across all cell populations, two cases (Cases 3 and 4) showed clear multiclonality in spatially distinct cell populations. In Case 3 (Fig. 3A and Supplementary Fig. S4A), 5 of 6 profiled cell populations (four CAH and one invasive endometrial carcinoma; from anterior and posterior aspects of the uterus) shared driving *PTEN*, *PIK3CA*, and *FBXW7* mutations; however, a population of invasive endometrial carcinoma (UT-25) lacked these alterations but harbored two distinct driving *PTEN* mutations and a *PIK3R1* mutation. In Case 4 (Fig. 3B), we profiled four regions of CAH from the uterine anterior, posterior, and fundus. Of note, although the CAH populations from the anterior and posterior aspect (UT-28 and UT-31) harbored the same driving *PTEN* and *KRAS* mutations, the two CAH populations from the fundus (UT-29 and UT-30) harbored distinct driving *PTEN* and *PIK3CA* mutations. Taken together, even with sampling of only an average of 5 spatially distinct cell populations per case, our results demonstrate that true multiclonality is relatively frequent during LGEC development.

Marked intratumoral heterogeneity in presumed LGEC driving mutations

Beyond multiclonality, we observed marked heterogeneity in presumed driving somatic mutations across LGEC and precursor populations in 10 of 13 cases, particularly those with both precursor and invasive endometrial carcinoma populations. For example, in Case 13 (Fig. 3C), we profiled separate populations of CAH (UT-87) and endometrial carcinoma (UT-83, 86, 87) from the posterior aspect. All populations shared driving *PTEN* (two mutations) and *KRAS* mutations. Although all endometrial carcinoma populations in this case also shared *PIK3R1* mutations, the CAH population harbored a distinct *PIK3R1* mutation. Likewise, a separate *PIK3R1* mutation was only present in two of the three endometrial carcinoma populations.

Similarly, in Case 2 (Fig. 3D and Supplementary Fig. S4B), we profiled separate populations of CAH (UT-19, anterior; UT-20, posterior) and ECsq (UT-21 and UT-22; fundus), all of which shared a driving *PTEN* mutation. Both ECsq populations also shared driving *CTNNB1* and *PIK3R1* mutations, neither of which was present in the CAH populations. However, both CAH populations shared a different driving *PIK3R1* mutation, whereas one CAH population harbored an additional missense *PIK3R1* mutation of unclear pathogenicity not present in the other CAH or ECsq populations. Case 11 showed similar intratumoral heterogeneity and branched evolution between precursor and invasive endometrial carcinoma populations as described in the Supplementary Results and Supplementary Fig. S3.

Intratumoral heterogeneity in candidate prognostic *CTNNB1* mutations

As described above, a motivator of this study was our previous observations of discordance in driving *CTNNB1* mutational status in two different endometrial carcinoma cases. Seven of 13 cases in our cohort showed no *CTNNB1* mutations in any cell population (clonally absent). In the remaining 6 cases where at least one population harbored a *CTNNB1* mutation, only one showed

Published OnlineFirst January 4, 2019; DOI: 10.1158/1541-7786.MCR-18-1178

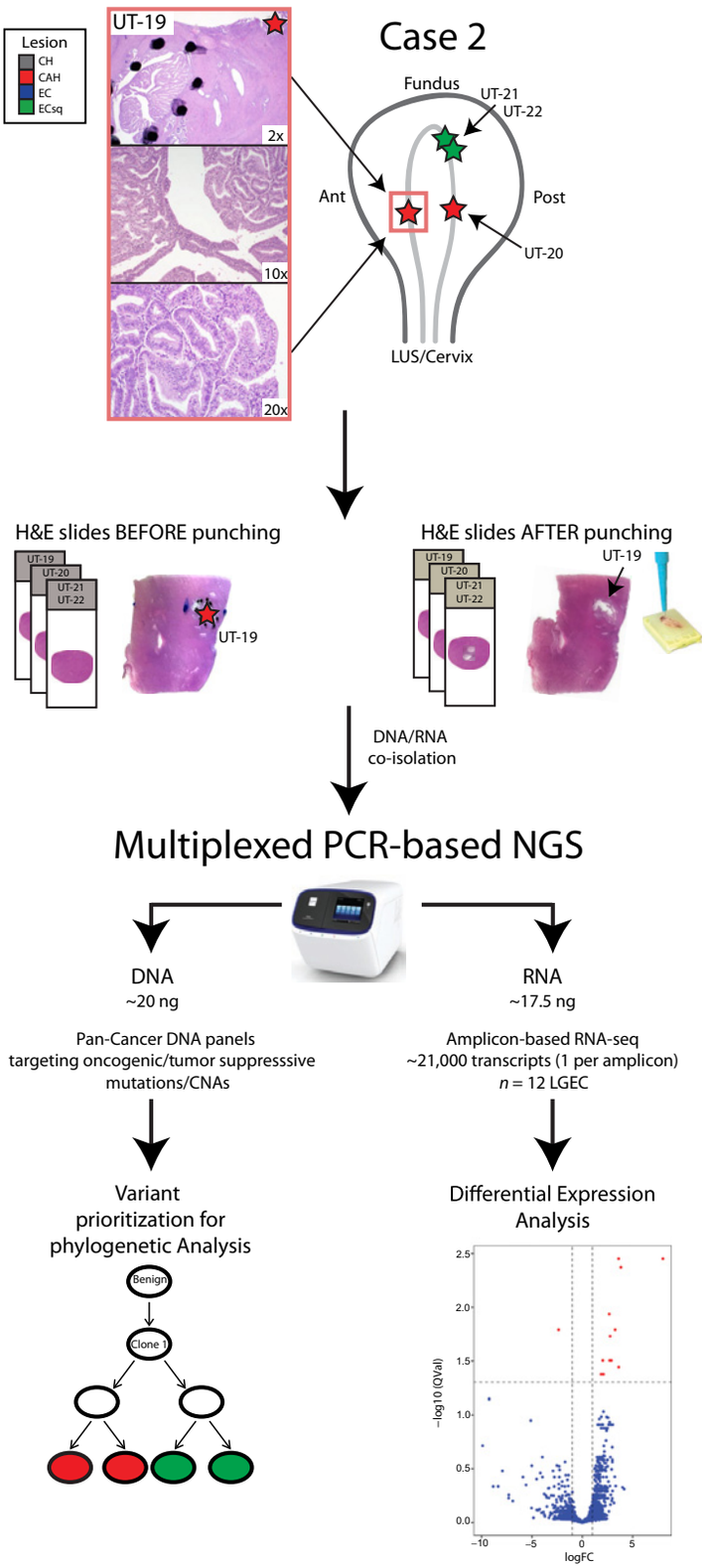

Figure 1.

Comprehensive DNA and RNA profiling of LGEC development from routine clinical specimens. Schematic of spatially defined uterine cell populations from a representative case (Case 2) is shown, with population type and associated histologic type indicated by the color scale (endometrial carcinoma, frankly invasive LGEC; sq, squamous metaplasia). Histology for the one population (UT-19, CAH) is shown with original magnification indicated. Precise tissue punching was used for isolation from routine FFPE blocks, and subsequent H&E-stained slides were used to confirm isolation of expected populations. Multiplexed PCR-based DNA and RNA sequencing was performed on $\leq$20 ng coisolated nucleic acids to comprehensively characterize LGEC development and intratumoral heterogeneity through driver gene alteration assessment and whole-transcriptome profiling.

Published OnlineFirst January 4, 2019; DOI: 10.1158/1541-7786.MCR-18-1178

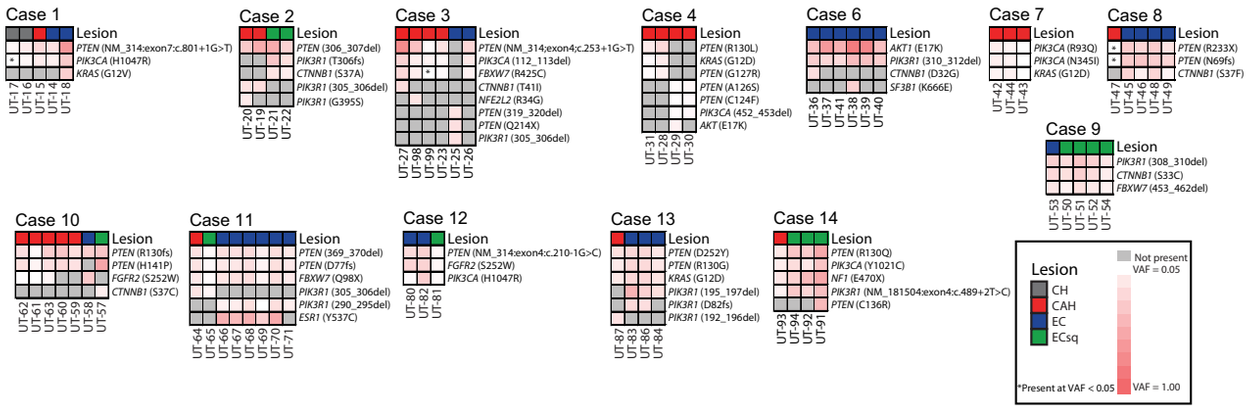

Figure 2.

Somatic mutations across LGEC development. Heatmaps showing all prioritized somatic mutations identified in endometrial cell populations, per LGEC case, with lesion histology indicated in the top row (according to the color scale at the bottom right). Individual somatic mutations are shown in rows, with the VAF indicated by the color hue gradient at the bottom right (gray, not present; *, well-supported reads on manual review and considered present but VAF < 5%).

clonal *CTNNB1* mutations in all profiled populations (Case 9, with endometrial carcinoma and ECsq populations). In Case 8, a shared *CTNNB1* mutation in all endometrial carcinoma populations was not present in the CAH sample (however, low tumor content in this sample precluded definitive exclusion). The remaining four cases showed: (1) a private (present in only one population) *CTNNB1* mutation in one precursor population but not in other precursor or endometrial carcinoma populations (Case 3), (2) shared *CTNNB1* mutations in all EC/ECsq populations but not in precursor populations (Case 2), (3) private *CTNNB1* mutation in only one of six endometrial carcinoma populations (Case 6), and (4) private *CTNNB1* mutation in one ECsq population but not in the endometrial carcinoma or multiple precursor populations (Case 10; Fig. 2). All mutations were observed at essentially clonal variant allele frequency (VAF), and mutational presence/absence was confirmed by Sanger sequencing (Supplementary Table S1 and Supplementary Fig. S5). Taken together, these results demonstrate the existence of intratumoral heterogeneity in potentially prognostic *CTNNB1* mutations both (1) within precursor and endometrial carcinoma populations and (2) within endometrial carcinoma populations in a given case.

Clonal sweep of heterogeneous mutations is common in LGEC

Heterogeneous mutations (those present in not all cell populations from a given case) may represent (1) subclonal alterations present but variably detected in all populations due to sampling or (2) clonal alterations present and selected for in the population. Through assessment of the VAF (# variant containing reads/total # reads) of truncal PI3K pathway driving mutations which inform on the estimated tumor content (VAF $\sim {}^{1}/_{2}$ and $\sim$ equivalent to the tumor content for heterozygous and homozygous variants, respectively), essentially all of the homogeneous (Supplementary Table S5) and heterogeneous (Supplementary Table S6) driving mutations observed in our cohort were present in all cells in the population [clonal cancer cell fraction (CCF)]. These results are consistent with selection of the variants due to increased fitness and "sweep" through the tumor cell population (22). Of note, the only gene that frequently showed less than clonal CCF was *CTNNB1* (Supplementary Table S6), further complicating its potential use as a prognostic and/or predictive biomarker. Importantly, however, the presence of numerous heterogeneous or private driving mutations in both precursor and endometrial carcinoma populations, most at clonal CCF, indicates a fitness advantage to these mutations regardless of histologic appearance or spatial proximity. Phylogenetic analysis thus supports extensive branched evolution in both precursor and endometrial carcinoma populations, consistent with branched evolution in LGEC development, in agreement with mechanisms in most other profiled cancers (23, 24).

Confirmation of intratumoral heterogeneity in CTNNB1 mutation–driven pathway activation through transcriptome sequencing

We next sought to further confirm the relevance of the often heterogeneous *CTNNB1* mutations by looking for transcriptional evidence of Wnt pathway activity in populations with and without *CTNNB1* mutations. Given the challenges of performing conventional or capture-based whole-transcriptome RNAseq with minute quantities of FFPE-isolated RNA (25), we attempted mxRNA-seq using $\leq$20 ng of coisolated RNA from 12 cell populations. Samples were selected to represent the spectrum of precursor versus endometrial carcinoma lesions with and without *CTNNB1* mutations. Across the 12 sequenced populations (with technical replicates in different batches), we generated an average of 8,891,762 end-to-end reads (Supplementary Table S7) with the 10,882 transcripts across the cohort showing >5 RPM used for further analysis. Technical replicates showed highly correlated normalized expression (median per pair Pearson $r = 0.9$; range, 0.92–0.99) with principal components analysis showing expected clustering of technical replicates (Fig. 4A). Differential expression analysis between profiled endometrial carcinoma ($n = 6$) and ECsq ($n = 2$) populations identified 18 transcripts overexpressed in ECsq versus endometrial carcinoma (Supplementary Fig. S6 and Supplementary Table S8). Comparison with transcripts overexpressed in TCGA lung squamous cell carcinoma versus adenocarcinoma (26) confirmed significant enrichment (OR, 66.9, two-sided Fisher exact test, $P = 1.15E$), including squamous epithelial-specific transcripts *PRR9*, *KRT31*, and *CALM3* (Supplementary Methods, Supplementary Results, and Supplementary

Published OnlineFirst January 4, 2019; DOI: 10.1158/1541-7786.MCR-18-1178

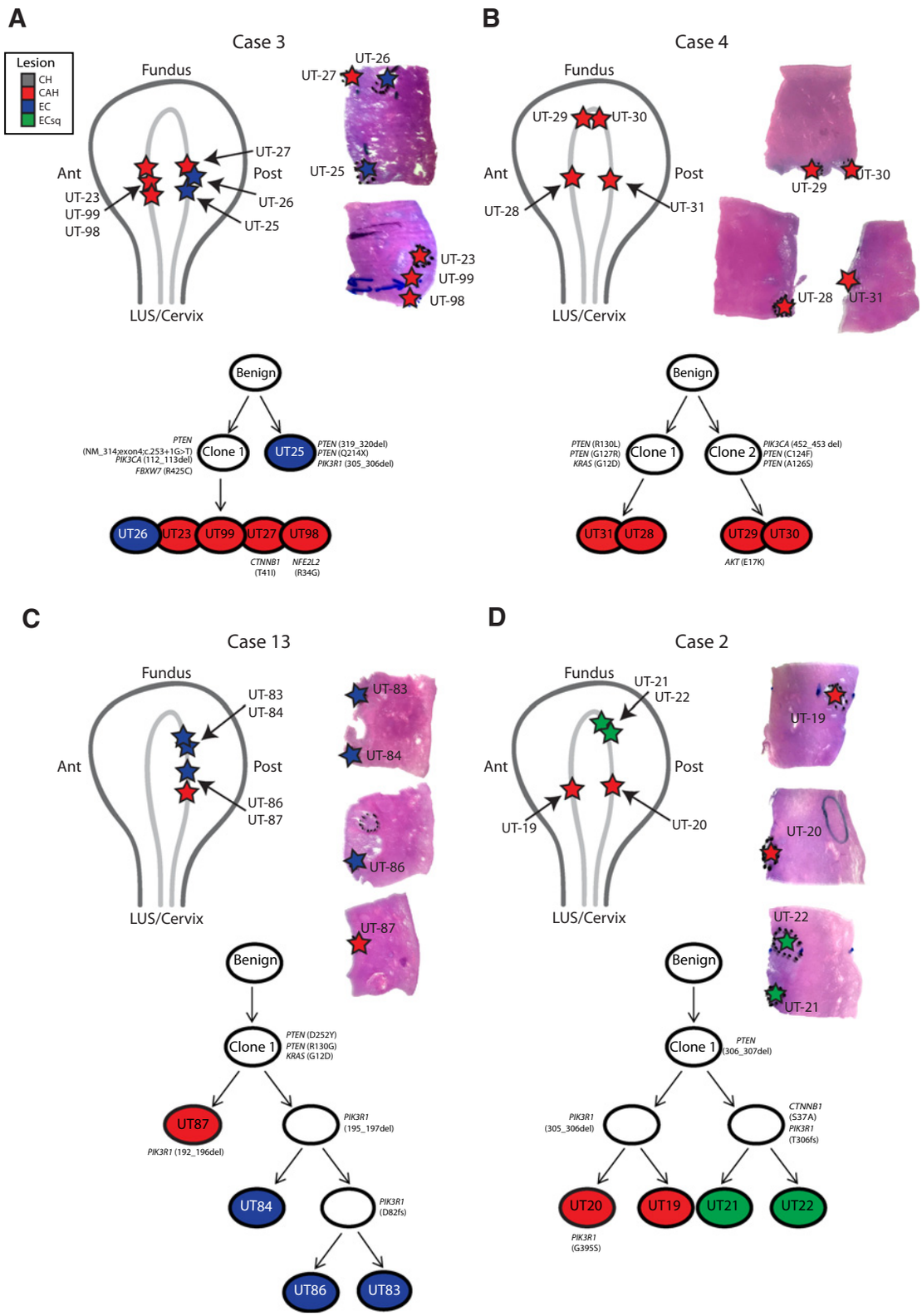

Figure 3.
Multifocality and marked intratumoral heterogeneity in LGEC development. For indicated cases, location of isolated cell populations (indicated by stars) in the uterus is indicated on anatomic diagrams and corresponding H&E slides. Phylogenetic trees for these cases are shown, with shared mutations for each clone indicated. **A** and **B,** Cases showing multiclonal LGEC development. **C** and **D,** Cases showing marked intratumoral heterogeneity in LGEC precursor and invasive cell populations.

Published OnlineFirst January 4, 2019; DOI: 10.1158/1541-7786.MCR-18-1178

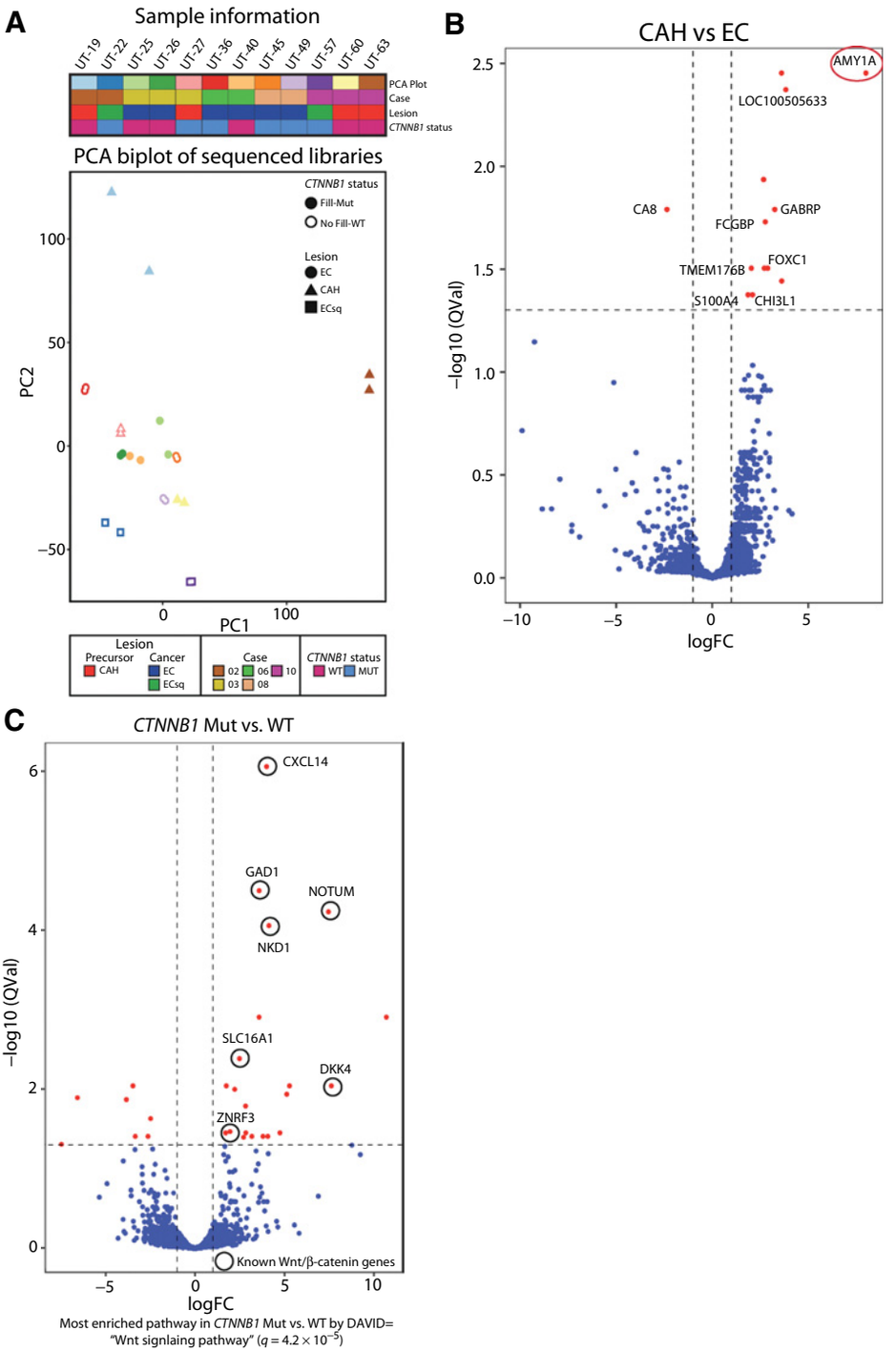

Figure 4.
Whole-transcriptome sequencing confirms deregulation of *Wnt* signaling in *CTNNB1*-mutated LGEC precursor and invasive populations. For the indicated LGEC cell populations with attributes indicated in the heatmap on top according to the legend, whole-transcriptome amplicon-based RNAseq was performed in duplicate from coisolated FFPE RNA. **A,** Principal component analysis (PCA) biplot of all sequenced samples, with samples colored according to the heatmap (PCA plot), and mutation status (filled vs. empty) and lesion type (shape) indicated. **B,** Volcano plot visualizing differentially expressed genes (FDR q value < 0.05) between precursor ($n = 4$) and invasive LGEC ($n = 8$) with genes of interest labeled. **C,** As in **B**, but showing differentially expressed genes in *CTNNB1*-mutant (MT, $n = 6$) vs. wild-type (WT, $n = 6$) populations, with canonical Wnt/β-catenin pathway genes circled and labeled.

Fig. S6), further supporting the validity of our approach. Likewise, we observed marked overexpression of *AMY1A* in profiled CAH ($n = 4$) vs. EC/ECsq ($n = 8$) populations and confirmed AMY1A protein overexpression in these samples by HC (Fig. 4B; Supplementary Figs. S7 and S8; Supplementary Table S8).

We thus assessed differentially expressed transcripts in profiled *CTNNB1*-mutant ($n = 6$) versus wild-type ($n = 6$) precursor and endometrial carcinoma populations. Importantly, we identified 21 transcripts significantly overexpressed in *CTNNB1*-mutant samples, including the known canonical Wnt target genes *NOTUM* (27), *CXCL14* (28), *GAD1* (29), *DKK4* (30), and *NKD1* (ref. 31; Fig. 4C and Supplementary Table S8). Database for Annotation Visualization and Integrated Discovery (DAVID) functional annotation assessment (32) also identified "Wnt signaling pathway" as the most significantly enriched biological process in the overexpressed *CTNNB1*-mutant gene set ($P =$

Published OnlineFirst January 4, 2019; DOI: 10.1158/1541-7786.MCR-18-1178

5.2×10^{-7}, Benjamini-corrected q value $= 4.2 \times 10^{-5}$). Gene set enrichment analysis of the hallmark gene sets also confirmed enrichment of the Wnt-β catenin signaling pathway [FDR q value = 0.032 (NES, 2.08); Supplementary Fig. S9]. Taken together, these results support the applicability of transcriptome-wide mxRNAseq to minute FFPE-isolated cell populations and confirm the functional relevance of shared and private *CTNNB1* mutations in both precursor and endometrial carcinoma populations.

Discussion

Here, to better understand the development of LGEC, we performed multiregion, comprehensive somatic molecular profiling of minute cell populations ranging from presumed precursor lesions through invasive LGEC. Through this high-depth (average >1,000x coverage) approach on spatially defined populations with variable histology from 13 cases, we identified marked intratumoral mutational heterogeneity in presumed cancer driving genes in the vast majority of cases. Our work builds on two small series of LGEC precursors and invasive components, which support substantial intratumoral heterogeneity and branched evolution in LGEC development (7, 8); however, our study is the first to definitively demonstrate multiclonality in both spatially defined precursor and invasive populations. Our findings have important implications for understanding LGEC development, as well as efforts to identify prognostic and predictive biomarkers, such as *CTNNB1*.

Consistent with the known role of PI(3)K pathway deregulation in EIN/LGEC development, all profiled cell populations harbored clear driving *PTEN*, *PIK3CA*, or *PIK3R1* mutations. In three cases, all populations harbored only *PIK3CA* or *PIK3R1* mutations, demonstrating that LGEC development does not absolutely require a *PTEN* mutation, consistent with TCGA data (1). Likewise, in 7 of 10 cases with driving clonal *PTEN* mutations in all populations, we only observed *PTEN* point mutations, in-frame short deletions, or splice site mutations, consistent with the lack of sensitivity of *PTEN* immunohistochemistry for EIN identification in pathologic practice (3, 33). Importantly, in several cases, we observed continued selection for, and convergent evolution in, driving mutations in the above three PI(3)K pathway members and *AKT1* in histologically presumed precursor lesions.

Through analysis of driving PI(3)K pathway mutations, we identified a case that developed two clonally distinct, multifocal LGECs (Case 3) and a case with clonally distinct precursor populations (Case 4). To our knowledge, such multiclonality has not been previously described in spatially defined populations. Remarkably, in Case 3, the two clonally distinct endometrial carcinoma populations were on the same FFPE block (<2 cm away), with no clear morphologic distinction between the invasive endometrial carcinoma populations (Fig. 3 and Supplementary Fig. S4A). In Case 4, where we could only sample superficial CAH-appearing populations, distinct clones were present in the uterine fundus versus anterior/posterior aspects. Given the relatively limited sampling performed in our study, we expect the observed rate of 15% multiclonality to be an underestimate, with more women developing multiple transformed LGEC populations.

As described above, the two cases showing discordant *CTNNB1* mutations in paired endometrial carcinoma samples (ref. 11 and C.S. Carter; in preparation) partly motivated this study. Importantly, Kurnit and colleagues and Liu and colleagues recently described *CTNNB1* mutations as prognostic in low-stage LGEC (9, 10). Similarly, a phase II clinical trial of everolimus and letrozole (NCT01068249) in women with endometrial carcinoma found particularly high response rates in those with endometrioid histology and *CTNNB1* mutations (34). Remarkably, in our LGEC cohort profiled herein, we observed clonal *CTNNB1* in only one of five cases (where at least one population harbored a *CTNNB1* mutation and tumor content was sufficiently high in all samples to enable confident assessment). In the remaining cases, we saw diverse intratumoral heterogeneity, with *CTNNB1* mutations being observed privately in precursor and not endometrial carcinoma populations (Case 3), shared in endometrial carcinoma but not precursor populations (Case 2), and private in endometrial carcinoma populations (Cases 6 and 10). Mutations in *CTNNB1* were also frequently subclonal in a given cell population, in contrast to essentially all other homogeneous or heterogeneous mutations in LGEC driver genes observed herein. Taken together, given that trials assessing *CTNNB1*—as well as PI3K members—as correlative biomarkers in women with endometrial carcinoma are ongoing (e.g., NCT02228681), our results suggest that sampling and assessment strategies have the potential to substantially affect results and should therefore be carefully considered during future trial design.

In addition to comprehensive DNA-based profiling, we also conducted amplicon-based whole-transcriptome sequencing on a subset of samples both to validate the approach, as well as determine whether subclonal *CTNNB1* mutations show evidence of transcriptional activation. Importantly, to our knowledge, this amplicon-based whole-transcriptome sequencing, which has the advantage of requiring <20 ng RNA, has only been reported in a single study of FFPE tissue (35). In addition to high pairwise concordance in technical replicates supporting the validity of our transcriptome data, we also confirmed expected differential transcript expression in ECsq versus endometrial carcinoma populations (overexpression of squamous epithelial transcripts) and *CTNNB1*-mutant versus wild type populations (Wnt/β-catenin target genes). In an exploratory analysis of precursor versus invasive endometrial carcinoma populations, we identified amylase (*AMY1A*) as markedly overexpressed in precursor populations and confirmed these findings in the same samples by immunohistochemistry using two anti-AMY1A antibodies. By IHC, amylase has been reported as only occasionally expressed in both benign secretory phase endometrial glands and well-differentiated endometrial adenocarcinomas (36–38), and hence we hypothesize that differential expression of AMY1A in paired precursor versus invasive populations likely reflects differentiation (AMY1A expression was not diffusely present in individual precursor appearing glands across individual sections or cases as shown in Supplementary Fig. S8) rather than a driving event in invasive endometrial carcinoma development. Importantly, through numerous lines of validation, our results demonstrate the applicability of amplicon-based whole-transcriptome sequencing to minute cell populations isolated from routine FFPE specimens, which may be particularly useful in scalable profiling of precursor lesions.

One of the major limitations of our study, which confounds most efforts to understand cancer development through precursors, is the use of concurrent presumed precursor and invasive populations to understand molecular features and drivers of invasive disease development. However, a more informative

Published OnlineFirst January 4, 2019; DOI: 10.1158/1541-7786.MCR-18-1178

study design, where precursor populations with or without subsequent development of invasive disease are compared, is confounded by the lack of clinical scenarios where precursor lesions are followed rather than completely excised. Our results herein combined with other studies highlight additional confounders including the potential of invasive disease to histologically mimic *in situ* precursors (39), the presence of multiclonal precursor and/or invasive clones, and the continued selection for driving mutations in precursor populations. Lastly, until studies profiling cases with confirmed nonprogressing precursor lesions are profiled, it is unclear whether such intratumoral heterogeneity in LGEC precursors and/or invasive populations is ubiquitous, or a more specific feature of "aggressive" behavior that may be exploited for diagnosis (as has been shown feasible in uterine aspirates; ref. 8) or prognosis (40). Supporting this concept, co-occurring *PTEN* and *PIK3CA* mutations have been reported to be extremely rare in CAH versus endometrial carcinoma (41). However, consistent with another study assessing *PIK3CA* and *PTEN* mutation frequency in CAH from cases with co-occurring endometrial carcinoma (42), we found co-occurrence of *PTEN* and *PIK3CA*/*PIK3R1* in at least one cell population with CAH histology in 6 of 8 cases (with *PTEN* mutations) with a co-occurring LGEC.

In summary, through comprehensive DNA and RNA profiling of minute, spatially defined populations from routine FFPE specimens, we demonstrate marked intratumoral heterogeneity and branched evolution in LGEC and precursors. Importantly, given this heterogeneity, sampling and sequencing depth may profoundly affect the detection of biomarkers in LGEC, including those such as *CTNNB1* that have been identified as prognostic in previous studies. Likewise, biomarker-based studies (such as those targeting PI3K pathway members) may also need to account for this heterogeneity. In addition, we also show relatively frequent true multiclonality, both of which have important implications for understanding LGEC development, predicting the behavior of LGEC precursors, and precision medicine for advanced LGEC. More generally, our approaches are applicable to archived FFPE samples and thus highly scalable, which may enable widespread sample contribution to efforts such as the PCGA (4), with the potential to transform the understanding of cancer precursors and early stage disease.

Disclosure of Potential Conflicts of Interest

S.A. Tomlins is Co-Founder and CMO at Strata Oncology and has provided expert testimony for Thermo Fisher Scientific. D.H. Hovelson is a Bioinformatician at Strata Oncology, is a consultant/advisory board member for Terumo BCT, and has provided expert testimony for Thermo Fisher Scientific. No potential conflicts of interest were disclosed by the other authors.

Authors' Contributions

Conception and design: L. Lazo de la Vega, K.R. Cho, S.A. Tomlins
Development of methodology: L. Lazo de la Vega
Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): L. Lazo de la Vega, M.C. Samaha, J. Siddiqui, A.P. Sciallis
Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): L. Lazo de la Vega, M.C. Samaha, K. Hu, D.H. Hovelson, K.R. Cho, A.P. Sciallis
Writing, review, and/or revision of the manuscript: L. Lazo de la Vega, M.C. Samaha, K. Hu, N.R. Bick, J. Siddiqui, C.S. Carter, K.R. Cho, S.A. Tomlins
Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): L. Lazo de la Vega, N.R. Bick, J. Siddiqui, C.-J. Liu
Study supervision: L. Lazo de la Vega, S.A. Tomlins

Acknowledgments

S.A. Tomlins was supported by the A. Alfred Taubman Medical Research Institute.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received November 1, 2018; revised December 2, 2018; accepted December 21, 2018; published first January 4, 2019.

References

1. Cancer Genome Atlas Research Network, Kandoth C, Schultz N, Cherniack AD, Akbani R, Liu Y, et al. Integrated genomic characterization of endometrial carcinoma. Nature 2013;497:67–73.
2. Kurman RJ. WHO classification of tumours of female reproductive organs. Lyon: International Agency for Research on Cancer; 2014. p. 307.
3. Sanderson PA, Critchley HO, Williams AR, Arends MJ, Saunders PT. New concepts for an old problem: the diagnosis of endometrial hyperplasia. Hum Reprod Update 2017;23:232–54.
4. Campbell JD, Mazzilli SA, Reid ME, Dhillon SS, Platero S, Beane J, et al. The case for a Pre-Cancer Genome Atlas (PCGA). Cancer Prev Res 2016;9:119–24.
5. Levine RL, Cargile CB, Blazes MS, van Rees B, Kurman RJ, Ellenson LH. PTEN mutations and microsatellite instability in complex atypical hyperplasia, a precursor lesion to uterine endometrioid carcinoma. Cancer Res 1998;58:3254–8.
6. van Esterik M, Van Gool IC, de Kroon CD, Nout RA, Creutzberg CL, Smit V, et al. Limited impact of intratumour heterogeneity on molecular risk assignment in endometrial cancer. Oncotarget 2017;8:25542–51.
7. Russo M, Broach J, Sheldon K, Houser KR, Liu DJ, Kesterson J, et al. Clonal evolution in paired endometrial intraepithelial neoplasia/atypical hyperplasia and endometrioid adenocarcinoma. Hum Pathol 2017;67:69–77.
8. Mota A, Colas E, Garcia-Sanz P, Campoy I, Rojo-Sebastian A, Gatius S, et al. Genetic analysis of uterine aspirates improves the diagnostic value and captures the intra-tumor heterogeneity of endometrial cancers. Mod Pathol 2017;30:134–45.
9. Kurnit KC, Kim GN, Fellman BM, Urbauer DL, Mills GB, Zhang W, et al. CTNNB1 (beta-catenin) mutation identifies low grade, early stage endometrial cancer patients at increased risk of recurrence. Mod Pathol 2017;30:1032–41.
10. Liu Y, Patel L, Mills GB, Lu KH, Sood AK, Ding L, et al. Clinical significance of CTNNB1 mutation and Wnt pathway activation in endometrioid endometrial carcinoma. J Nat Cancer Inst 2014;106 pii:dju245.
11. McDaniel AS, Stall JN, Hovelson DH, Cani AK, Liu CJ, Tomlins SA, et al. Next-generation sequencing of tubal intraepithelial carcinomas. JAMA Oncol 2015;1:1128–32.
12. Hanauer DA, Mei Q, Law J, Khanna R, Zheng K. Supporting information retrieval from electronic health records: a report of University of Michigan's nine-year experience in developing and using the Electronic Medical Record Search Engine (EMERSE). J Biomed Inform 2015;55:290–300.
13. Hovelson DH, McDaniel AS, Cani AK, Johnson B, Rhodes K, Williams PD, et al. Development and validation of a scalable next-generation sequencing system for assessing relevant somatic variants in solid tumors. Neoplasia 2015;17:385–99.
14. Grasso C, Butler T, Rhodes K, Quist M, Neff TL, Moore S, et al. Assessing copy number alterations in targeted, amplicon-based next-generation sequencing data. J Mole Diagn 2015;17:53–63.
15. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 2010;26:139–40.

Published OnlineFirst January 4, 2019; DOI: 10.1158/1541-7786.MCR-18-1178

16. McCarthy DJ, Chen Y, Smyth GK. Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. Nucleic Acids Res 2012;40:4288–97.
17. Benjamini Y, Hochberg Y. Controlling the false discovery rate - a practical and powerful approach to multiple testing. J Roy Stat Soc B Met 1995;57: 289–300.
18. Uhlen M, Zhang C, Lee S, Sjostedt E, Fagerberg L, Bidkhori G, et al. A pathology atlas of the human cancer transcriptome. Science 2017;357 pii:eaan2507.
19. Oda K, Stokoe D, Taketani Y, McCormick F. High frequency of coexistent mutations of PIK3CA and PTEN genes in endometrial carcinoma. Cancer Res 2005;65:10669–73.
20. Urick ME, Rudd ML, Godwin AK, Sgroi D, Merino M, Bell DW. PIK3R1 (p85alpha) is somatically mutated at high frequency in primary endometrial cancer. Cancer Res 2011;71:4061–7.
21. Cheung LW, Hennessy BT, Li J, Yu S, Myers AP, Djordjevic B, et al. High frequency of PIK3R1 and PIK3R2 mutations in endometrial cancer elucidates a novel mechanism for regulation of PTEN protein stability. Cancer Discov 2011;1:170–85.
22. Sottoriva A, Kang H, Ma Z, Graham TA, Salomon MP, Zhao J, et al. A Big Bang model of human colorectal tumor growth. Nat Genet 2015;47: 209–16.
23. Sun R, Hu Z, Curtis C. Big Bang tumor growth and clonal evolution. Cold Spring Harb Perspect Med 2017;8 pii:a028381.
24. Hiley C, de Bruin EC, McGranahan N, Swanton C. Deciphering intratumor heterogeneity and temporal acquisition of driver events to refine precision medicine. Genome Biol 2014;15:453.
25. Cieslik M, Chugh R, Wu YM, Wu M, Brennan C, Lonigro R, et al. The use of exome capture RNA-seq for highly degraded RNA with application to clinical cancer sequencing. Genome Res 2015;25:1372–81.
26. Wang Q, Armenia J, Zhang C, Penson AV, Reznik E, Zhang L, et al. Unifying cancer and normal RNA sequencing data from different sources. Sci Data 2018;5:180061.
27. Kakugawa S, Langton PF, Zebisch M, Howell S, Chang TH, Liu Y, et al. Notum deacylates Wnt proteins to suppress signalling activity. Nature 2015;519:187–92.
28. Park BY, Hong CS, Sohail FA, Saint-Jeannet JP. Developmental expression and regulation of the chemokine CXCL14 in Xenopus. Int J Dev Biol 2009; 53:535–40.
29. Li CM, Kim CE, Margolin AA, Guo M, Zhu J, Mason JM, et al. CTNNB1 mutations and overexpression of Wnt/beta-catenin target genes in WT1-mutant Wilms' tumors. Am J Pathol 2004;165:1943–53.
30. Pendas-Franco N, Garcia JM, Pena C, Valle N, Palmer HG, Heinaniemi M, et al. DICKKOPF-4 is induced by TCF/beta-catenin and upregulated in human colon cancer, promotes tumour cell invasion and angiogenesis and is repressed by 1alpha,25-dihydroxyvitamin D3. Oncogene 2008;27: 4467–77.
31. Larraguibel J, Weiss AR, Pasula DJ, Dhaliwal RS, Kondra R, Van Raay TJ. Wnt ligand-dependent activation of the negative feedback regulator Nkd1. Mol Biol Cell 2015;26:2375–84.
32. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 2009; 4:44–57.
33. Pavlakis K, Messini I, Vrekoussis T, Panoskaltsis T, Chrissanthakis D, Yiannou P, et al. PTEN-loss and nuclear atypia of EIN in endometrial biopsies can predict the existence of a concurrent endometrial carcinoma. Gynecol Oncol 2010;119:516–9.
34. Slomovitz BM, Jiang Y, Yates MS, Soliman PT, Johnston T, Nowakowski M, et al. Phase II study of everolimus and letrozole in patients with recurrent endometrial carcinoma. J Clin Oncol 2015;33:930–6.
35. FitzGerald LM, Jung CH, Wong EM, Joo JE, Gould JA, Vasic V, et al. Obtaining high quality transcriptome data from formalin-fixed, paraffin-embedded diagnostic prostate tumor specimens. Lab Invest 2018; 98:537–50.
36. Ueda G, Yamasaki M, Inoue M, Tanaka Y, Inoue Y, Nishino T, et al. Immunohistochemical demonstration of amylase in endometrial carcinomas. Int J Gynecol Pathol 1986;5:47–51.
37. Ueda G, Yamasaki M, Inoue M, Tanaka Y, Inoue Y, Nishino T, et al. Capacity for amylase production of endometrial carcinomas. Nihon Sanka Fujinka Gakkai Zasshi 1985;37:305–6.
38. Lee YS, Raju GC. The expression and localization of amylase in normal and malignant glands of the endometrium and endocervix. J Pathol 1988;155: 201–5.
39. Haffner MC, Weier C, Xu MM, Vaghasia A, Gurel B, Gumuskaya B, et al. Molecular evidence that invasive adenocarcinoma can mimic prostatic intraepithelial neoplasia (PIN) and intraductal carcinoma through retrograde glandular colonization. J Pathol 2016;238:31–41.
40. Espiritu SMG, Liu LY, Rubanova Y, Bhandari V, Holgersen EM, Szyca LM, et al. The evolutionary landscape of localized prostate cancers drives clinical aggression. Cell 2018;173:1003–13 e15.
41. Hayes MP, Wang H, Espinal-Witter R, Douglas W, Solomon GJ, Baker SJ, et al. PIK3CA and PTEN mutations in uterine endometrioid carcinoma and complex atypical hyperplasia. Clin Cancer Res 2006;12(20 Pt 1): 5932–5.
42. Berg A, Hoivik EA, Mjos S, Holst F, Werner HM, Tangen IL, et al. Molecular profiling of endometrial carcinoma precursor, primary and metastatic lesions suggests different targets for treatment in obese compared to non-obese patients. Oncotarget 2015;6:1327–39.

Published OnlineFirst January 4, 2019; DOI: 10.1158/1541-7786.MCR-18-1178

Molecular Cancer Research

Multiclonality and Marked Branched Evolution of Low-Grade Endometrioid Endometrial Carcinoma

Lorena Lazo de la Vega, Mia C. Samaha, Kevin Hu, et al.

Mol Cancer Res 2019;17:731-740. Published OnlineFirst January 4, 2019.